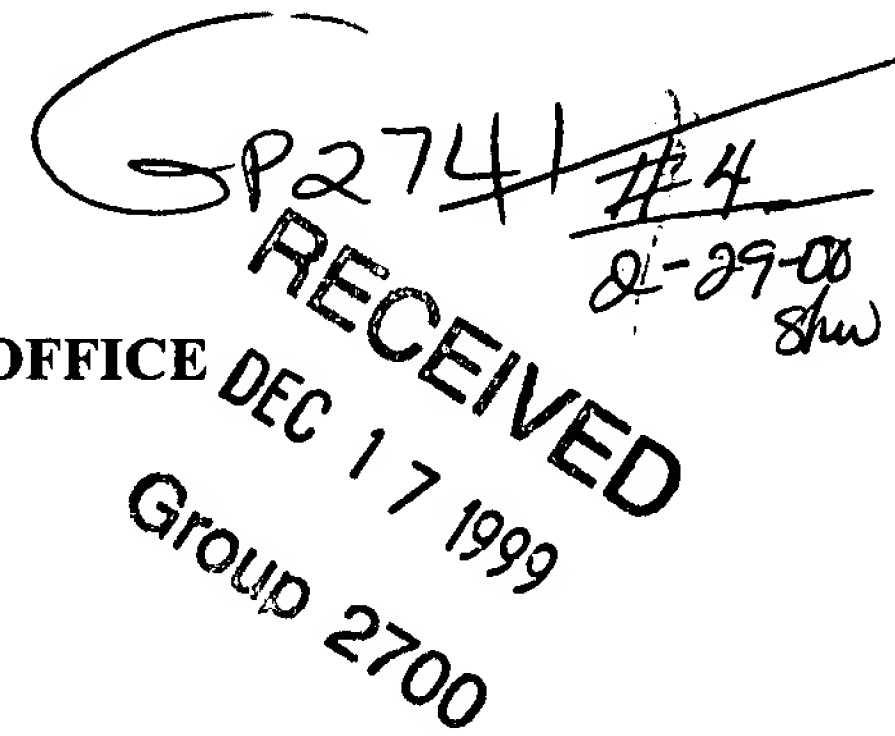
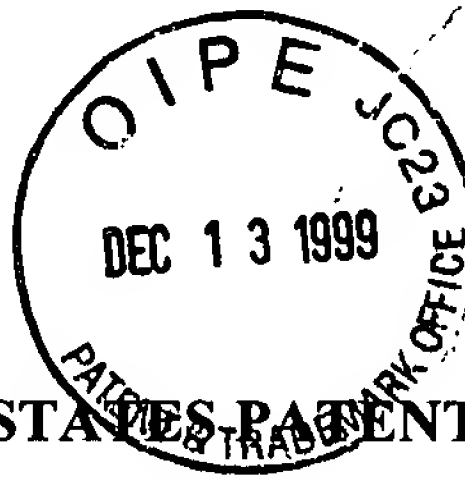


Docket No.: AND1P111



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Valery A Pertrushin

App. Ref.: AND1P111

Examiner: Not Assigned

Serial No.: 09/388,909

Filing Date: 8/31/99

Art Unit: 2741

Title:

SYSTEM, METHOD, AND
ARTICLE OF MANUFACTURE FOR
DETECTING EMOTION IN VOICE
SIGNALS BY UTILIZING STATISTICS
FOR VOICE SIGNAL PARAMETERS

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as First Class Mail in an envelope addressed to: Assistant Commissioner for Patents, Washington, DC 20231 on December 9, 1999.

Signed: 

Julie A. Curts

Assistant Commissioner for Patents
Washington D.C. 20231

PETITION TO MAKE SPECIAL
37 C.F.R. 1.102 and MPEP § 708.02(VIII)

Sir:

1. Petition

Applicant hereby petitions to make this new application special. This application has not received any examination by the Examiner.

2. Fee

The Office is authorized to charge the required fee for this petition to deposit account **50-0797**, of Andersen Consulting, LLP. At any time during the pendency of this application, please charge any fees required or credit any overpayments to the aforementioned deposit account. A duplicate copy of this petition (cover and signature pages only) is enclosed for billing purposes.

3. Claims

All of the claims in this case are directed to a single invention. If the Office determines that all of the claims presented are not directed to a single invention, then applicant will make an election without traverse as a prerequisite to the grant of special status.

4. Search

A preliminary patentability search was performed by a technical expert within our firm in databases of U.S. Patents in the following fields: 704/270 and 704/275 for a system, method and article of manufacture for detecting emotion using statistics. To accomplish this, a database is provided. The database has statistics including human associations of voice parameters with emotions. Next, a voice signal is received. At least one feature is extracted from the voice signal. Then the extracted voice feature is compared to the voice parameters in the database. An emotion is selected from the database based on the comparison of the extracted voice feature to the voice parameters and is then output. Particular keywords used include: "emotion", "voice", "statistics", and "voice parameter." This and related searches revealed 19 references, each of which is discussed in the petition.

5. Discussion of Related References

There is submitted herewith a copy of each of the references deemed most closely related to the subject matter of the claimed invention. Also attached is Form PTO-1449.

United States Patent Number 4,490,840 to Jones

A method for analyzing vocal sounds of organisms, particularly humans, for characteristics defined as voice-style (resonance, quality), speech-style (variable-monotone, choppy-smooth, etc.), and perceptual-style (sensory-internal, hate-love, etc.). The amount of each characteristic is calculated from relative and difference values of measured elements including six spectral peaks and pauses. Coefficient tables indicate the relative contribution of measured elements.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions, where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. More particularly, Jones does not disclose or suggest utilization of statistics for detecting emotion.

United States Patent Number 5,909,665 to Kato

A speech recognition system that includes an analyzing unit for extracting a sound, sequentially dividing the sound into a plurality of frames, converting each of the frames sequentially to first data, and sequentially storing the first data to an input pattern memory, a distance calculating unit for reading a predetermined number of the first data from the input pattern memory, reading one of second data from a standard pattern memory, calculating first distances between each of the predetermined number of the first data and the one of the second data, and a judging unit for judging a word representing the sound based on the first distances.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. In particular, Kato teaches using distances rather than statistics to analyze speech.

United States Patent Number 4,592,086 to Watari, et al.

A continuous speech recognition system that determines the similarity between input patterns and reference patterns over time such that similarities between previously spoken speech patterns and reference patterns are determined while speech continues to be spoken. Degrees of dissimilarity at arbitrary reference pattern word times are determined asymptotically and are recorded. The minimum degree of dissimilarity is determined and the corresponding word is categorized. Recognition decisions are ultimately made in reverse chronological order.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. Particularly, Watari et al. fails to teach or suggest detecting an emotion in the speech.

United States Patent Number 5,539,861 to DeSimone

A method for improving the recognition rate of a speech recognition system by compensating for changes in the user's speech that result from factors such as emotion, anxiety or fatigue. A speech signal derived from a user's utterance is modified by a preprocessor and provided to a speech recognition system to improve the recognition rate. The speech signal is modified based on a bio-signal which is indicative of the user's emotional state.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. More particularly, DeSimone teaches using bio-signals to detect emotion rather than extracted voice parameters.

United States Patent Number 4,142,067 to Williamson

A speech analyzer for determining the emotional state of a person by analyzing pitch or frequency perturbations in the speech pattern. The analyzer determines null points or "flat" spots in a FM demodulated speech signal and it produces an output indicative of the nulls. The output can be analyzed by the operator of the device to determine the emotional state of the person whose speech pattern is being monitored.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. More specifically, Williamson does not utilize statistics when determining the emotional state of a person by analyzing the speech pattern.

United States Patent Number 4,093,821 to Williamson

A speech analyzer for determining the emotional state of a person by analyzing pitch or frequency perturbations in the speech pattern. The analyzer determines null points or "flat" spots in an FM demodulated speech signal and produces a first output indicative of the nulls and a second output indicative of the presence of a "word." A pitch frequency processor receives the FM demodulated speech signal and the first output of the detector means and produces an output having an amplitude proportional to the frequency of the speech signal at the null. A pitch null duration processor receives the first output of the detector means and produces an output having an amplitude proportional to the duration of the nulls. A ratio processor receives the first and second outputs of the detector means and produces an output proportional to the ratio of the total duration of all the nulls within a word to the total duration of the word. The outputs of the pitch frequency processor, pitch null duration processor and ratio processor can be used to provide an indication of the emotional state of the individual whose speech is being analyzed.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. In particular, Williamson does not utilize statistics when determining the emotional state of a person by analyzing the speech pattern.

United States Patent Number 3,971,034 to Bell, Jr., et al.

A method of detecting psychological stress by evaluating manifestations of physiological change in the human voice wherein the utterances of a subject under examination are transduced to electrical signals and processed to emphasize selected characteristics which have been found to change with psycho-physiological state changes. The processed signals are then displayed, as on a strip chart recorder, for observation, comparison and analysis. An especially useful characteristic is an infrasonic modulation in the voice. Apparatus for performing detection of this type includes a transducer, a magnetic recorder, a series diode, a plurality of integrating capacitors, an amplifier and a chart recorder. A second apparatus includes filter means, an FM discriminator and a detector, a waveform integrator, an amplifier and a recorder for producing a visible record.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. More specifically, Bell, Jr. et al. determines psychological stress rather than emotion.

United States Patent Number 5,163,083 to Dowden, et al.

Methods and apparatus for automatically processing operator assistance calls. A caller is connected to an automated operator position. The automated position has speech recognition facilities to replace those of an operator, has announcement

capabilities to replace those of an operator, and has control apparatus for transmitting and receiving the same set of messages transmitted and received by an operator position. The operator assistance switch has the same interface to an automated position as to an operator position and interacts with the two identically. Since the capabilities of the automated position are limited by its program, the automated position switches a call to an automated position when a situation occurs for which it has not been programmed. A switch need not be specially programmed to communicate with an automated position. New operator assistance services can be provided automatically without rewriting the complex control software of the switch.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. More particularly, Dowden et al. teaches recognizing speech patterns rather than detecting emotion.

United States Patent Number 5,936,515 to Right et al.

A field programmable audible signal having voice message annunciating capability and a field programming device. The signal has two separate field programming paths. One path includes a built in microphone and the other is a facility to receive a download voice message from a field programming device as by a cable that plugs into both the signal and the programming device. The field programming device is capable of providing either of two messages during a download operation. The field programming device includes a record facility to change at least one of the messages and is small enough to fit in a hand held housing.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion

based on the comparison of the extracted voice feature to the voice parameters. In particular, Right et al. does not detect emotion from a voice signal.

United States Patent Number 4,996,704 to Brunson

An electronic messaging system that allows a system subscribed to record a plurality of "customized" announcement messages. Each such message is associated with at least one calling party. Upon receiving an incoming communication for that subscriber, the system automatically utilizes the calling party identification for that communication to retrieve the associated customized announcement message. The calling party identification, which identifies the communication instrument utilized by the calling party, is automatically provided to the electronic messaging system by the communications network through which the incoming communication is routed.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. More specifically, Brunson does not detect emotion from a voice signal.

United States Patent Number 5,495,553 to Jakatdar

A recognition device and method for recognizing a voice message in the form of pulse code modulation (PCM) digital signals indicative of samples of the voice message. The device and method are adapted such that a recognition result is not provided if the digital signal content satisfies certain requirements which are indicative of a likely erroneous recognition result. The recognition device and method are further adapted to reduce errors in recognizing voice messages of the same message content but different amplitude, as well as to permit simultaneous recognition and storing for recording of a voice message.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of

voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. More specifically, Jakatdar fails to detect emotion in the voice message.

United States Patent Number 4,696,038 to Doddington

A voice messaging system with an LPC analyzer in combination with a pitch extractor, wherein LPC parameters and a residual signal organized in a sequence of speech data frames are provided by the LPC analyzer as an output representative of an analog speech signal. The pitch extractor is operably associated with the LPC analyzer and produces a plurality of pitch candidates for each of the speech data frames in the sequence thereof. Dynamic programming is performed on the plurality of pitch candidates for each speech data frame and also with respect to a voiced/unvoiced decision of the speech data for each frame by tracking both pitch and voicing from frame to frame to provide an optimal pitch value and also an optimal voicing decision. During dynamic programming, a cumulative penalty for a sequence of frame pitch/voicing decisions is accumulated by defining a transition error between each pitch candidate of a current speech data frame and each pitch candidate of the preceding frame, and defining a cumulative error for each pitch candidate of the current frame equal to the transition error between the pitch candidate of the current frame plus the cumulative error of an optimally identified pitch candidate in the preceding frame to locate the track providing optimal pitch and voicing decisions based upon the lowest cumulative penalty. An encoder then encodes the LPC parameters as generated by the LPC analyzer and the optimal pitch and voicing decisions for each speech data frame for subsequent use in providing an audible synthesized speech output substantially identical to the original speech input.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion

based on the comparison of the extracted voice feature to the voice parameters. In general, Doddington fails to use statistics to detect emotion in a voice signal.

United States Patent Number 4,602,129 to Matthews et al.

An advanced electronic telecommunications system for the deposit, storage and delivery of audio messages to both users and non-users with limited access provided to the non-user under the control of the user. A Voice Message System interconnects multiple private exchanges of a subscriber with a central telephone office. Individual subscriber users may access the Voice Message System through ON NET telephones or OFF NET telephones. Selected non-users may be allowed access through the OFF NET telephones, the scope of the access of the selected non-users being determined by a subscriber user. The Voice Message System includes an administrative subsystem, call processor subsystem and a data storage subsystem. The Voice Message System enables the user to deposit a message in data storage subsystem for automatic delivery to other addresses connected to the system and to designate the message for priority transmission. The recipient is able to redirect the message from a message originator to a second recipient and the second recipient can re-redirect it to a third recipient. The Voice Message System also enables a user to access the system to determine if any messages have been in data storage subsystem for him. Prerecorded instructional messages are deposited in the data storage subsystem for instructing a user or a selected non-user on their progress in using the system.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. More specifically, Matthews et al. does not detect emotion from a voice signal.

United States Patent Number 5,913,196 to Talmor et al.

A system for establishing an identity of a speaker including a computerized system which includes at least two voice authentication algorithms. Each of the at least

two voice authentication algorithms is different from one another and serves for independently analyzing a voice of the speaker for obtaining an independent positive or negative authentication of the voice by each of the algorithms. If every one of the algorithms provide positive authentication, the speaker is positively identified, whereas, if at least one of the algorithms provides negative authentication, the speaker is negatively identified.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. In more detail, Talmor et al. teaches a system for identifying a speaker, not determining the speaker's emotion.

United States Patent Number 5,903,870 to Kaufman

A speech transducer, a processor, and a display device. The display device comprises a screen. The processor produces a plurality of windows on the screen at the same time, at least two of the windows comprised of different types of data. The processor also receives a speech signal from the speech transducer and modifies a parameter of one or more of the windows based on the speech signals. A plurality of data sources are provided at least two of which produce different types of data. Preferably one or more windows each comprising data from a different data source, are produced on the screen at the same time. The windows on the screen are arranged in a grid comprised of a plurality of rows and a plurality of columns. The processor includes a voice input device for translating speech electrical signals into language signals and a language device for implementing language signals to modify a window on the screen of the display device.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the

voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. More particularly, Kaufman fails to teach or suggest detection of the speaker's emotion.

United States Patent Number 5,812,977 to Douglas

A computer assisted system that enables a computer user with less than fully developed computer skills to enable and implement a number of subroutines. The disclosed system, which is preferably operated by means of voice commands, therefore improves the performance of the user so that the subroutines can be fetched more readily, operated more effectively to obtain the desired results or output, and then easily closed or terminated. The disclosed system further simplifies computer start up operations.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. More specifically, Douglas teaches a voice recognition system rather than one that detects emotion.

United States Patent Number 5,774,859 to Houser et al.

A system for controlling a device such as a television and for controlling access to broadcast information such as video, audio, and/or text information. The system includes a first receiver for receiving utterances of a speaker, a second receiver for receiving vocabulary data defining a vocabulary of utterances, and a processor for executing a speech recognition algorithm using the received vocabulary data to recognize the utterances of the speaker and for controlling the device and the access to the broadcast information in accordance with the recognized utterances of the speaker.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of

voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. More specifically, Houser et al. discloses a voice control system but does not teach or suggest detecting emotion in the voice.

United States Patent Number 5,884,247 to Christy

A method and apparatus for language translation by representing natural-language sentences in accordance with a constrained grammar and vocabulary structured to permit direct substitution of linguistic units in one language for corresponding linguistic units in another language. Preferably, the vocabulary is represented in a series of physically or logically distinct databases, each containing entries representing a form class as defined in the grammar. Translation involves direct lookup between the entries of a reference sentence and the corresponding entries in one or more target languages.

The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. In particular, Christy teaches a translating method rather than an emotion detector.

United States Patent Number 5,893,057 to Fujimoto, et al.

Speaker recognition methods and systems that involve at least two processing units for performing the speaker recognition based upon his or her voice input. To perform the speaker recognition efficiently as well as securely, the voice input is initially processed at the input site so that intermediate voice characteristic information is extracted. The intermediate voice characteristic information is transmitted to a second location for the final determination for identifying or verifying a speaker.



The patent fails to disclose, teach or suggest the system, method and article of manufacture for detecting emotion using statistics which include human associations of voice parameters with emotions where one or more features are extracted from the voice signal and compared to the voice parameters to select and output an emotion based on the comparison of the extracted voice feature to the voice parameters. More particularly, Fujimoto et al. teaches an identification system rather than emotion detection.

RECEIVED
DEC 17 1999
Group 2700

Thus, for the reasons stated above, the limitations of Applicant's claims 1, 7, and 13, are believed to be completely foreign to the teachings of the references cited herein and therefore are believed to be allowable over the cited references. Applicant's claims 2-6 depend from Applicant's claim 1, claims 8-12 depend from Applicant's claim 7, and claims 14-18 depend from Applicant's claim 13, and therefore, by virtue of their dependency are also believed to be allowable over the cited reference.

6. Declaration

As the undersigned practitioner, being duly registered to practice before the U.S. Patent and Trademark Office, I declare that I have made or caused to be made the careful and thorough search of the prior art as described herein.

Hickman Stephens & Coleman, LLP
P.O. Box 52037
Palo Alto, California 94303

Telephone: 408.558.9950
Facsimile: 408.558.9960

Respectfully Submitted,

Keith Stephens
Reg. No. 32,632